1971

Retroactive inhibition in free recall as a function of list organizations.

Jane Perlmutter
University of Massachusetts Amherst

RETROACTIVE INHIBITION IN FREE RECALL AS A FUNCTION OF LIST ORGANIZATION

A thesis Presented

By

Jane Perlmutter

Submitted to the Graduate School of the
University of Massachusetts in
partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

November (Month) 1971 (Year)

DEPARTMENT OF PSYCHOLOGY

RETROACTIVE INHIBITION IN FREE RECALL AS A FUNCTION OF LIST ORGANIZATION

A thesis Presented

By

Jane Perlmutter

Approved as to style and content by:

(Chairman of Committee)

(Head of Department)

(Member)

(Member)

November 1971
(Month) (Year)

ACKNOWLEDGEMENTS

The author wishes to thank the members of her thesis committee, Dr. James Chumbly and Dr. John Emrick, for their thoughtful suggestions and criticisms concerning this thesis. Special thanks are due Dr. James M. Royer for his invaluable assistance throughout the preparation of this thesis.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Retroactive Inhibition

Retroactive inhibition (RI) is the decrement in retention attributable to interpolated learning. The most common type of RI study is one in which a particular variable is manipulated in the acquisition phase of the experiment, and the loss of words from an initially learned list is examined as a function of the manipulation. The literature on RI has been reviewed a number of times in the last several decades (i.e., Swenson, 1941; Slamecka and Ceraso, 1960; and Keppel, 1968). Slamecka and Ceraso make use of the following classification for independent variables which have been investigated: 1)degree of acquisition; 2)similarity of materials; 3)extrinsic factors; and 4)temporal effects.

The dominant theoretical explanation of RI in the late 1930's and early 1940's was a competition theory advanced chiefly by McGeoch and his associates (see McGeoch and Irion, 1952). This theory attributed the failure to recall an old association to either: 1) the greater strength of the new associations; 2)a mutual blocking of old and new associations; or 3)a confusion between the two components. The first, which Barnes and Underwood (1959) called the independence hypothesis, asserts that the associative strength of the response in the first list is not changed

by learning a second list. The second hypothesis, response dominance at recall, specifies the mechanism responsible for the observed retention loss (i.e., the displacement of the desired response in one list by a stronger response from the other list).

McGeoch's position remained dominant until 1940 when Melton and Irwin (1940) proposed the first multi-factor theory of RI. In their study, Melton and Irwin (1940) noted that interlist intrusions and RI peaked separately as a function of interpolated learning (IL) trials (intrusions were maximal after 10 trials, while RI continued to increase until 20 IL trials). This led them to conclude that the RI attributable to competition was only a small portion of total RI observed. The remaining RI was attributed to an unexplained "factor X" (later known as unlearning). A similarity between the phenomenon of unlearning and the phenomenon of extinction in classical conditioning has been noted by several writers (i.e., Underwood, 1948a, 1948b; Melton, 1961; and Postman, 1961). The assumption was that unlearning occurs as a result of the nonreinforced elicitation of a first list response during second list learning. In addition, the responses which became unavailable can be easily relearned; a condition similar to spontaneous recovery. Keppel (1968) sum-

marized the current explanation of the two factors operating as follows:

> unlearning of the originally learned response (OL) as a consequence of interpolated learning (IL), and competition between OL and IL responses at the time of recall.

One difficulty in the research based on a two factor theory has been to separate the effects of unlearning from those attributable to competition. In the 1950's two techniques were introduced which attempted to separate response competition from unlearning. They were the modified free recall (MFR) technique, and the modified modified free recall (MMFR) technique. The MFR technique was used by Briggs (1954), and involved presenting the stimulus term and asking the S for the first response that comes to mind (OL or IL) periodically throughout OL and IL. This technique allowed Briggs to assess the relative strength of responses. He found that the number of OL responses given during IL decreased as IL trials increased. Although this technique seemed to separate competition and unlearning, the possibility remained that the Ss had responses available which they did not give. Barnes and Underwood (1959) reported a study which attempted to answer the question as to whether or not the first list responses were available for recall. The technique which Barnes and Underwood (1959) used, later became known as MMFR, and involved asking the S for both first and

second list responses to a given stimulus. They found that first list words became unavailable with increased IL trials.

Retroactive Inhibition in Free Recall

Much of the previous research has used the paired-associate learning paradigm as a vehicle for examining RI. More recently, free recall techniques have also been used. Free recall (FR) involves either simultaneous or successive presentation of items to Ss, followed by a recall session in which the order of recall is determined by the S. This technique has been used in both single-trial and multi-trial experiments.

One of the first studies of RI in FR was reported by Tulving and Thorton (1959). Their experiment included groups which learned 1, 2, 3, 4, or 5 lists and then were requested to recall all of the words from the lists they just studied. They found recall of the first list to be inversely related to the total number of lists which were learned.

Another study which used FR to study RI was reported by Postman and Keppel (1967). They stated that the question their experiment was trying to answer was "whether and to what extent RI in FR exhibits functional characteristics which parallel those observed in studies of unlearn-

ing with paired-associates." Their study paralleled a study by Barnes and Underwood (1959) which had Ss learn a first list of nonsense syllable-adjective pairs to a criterion of one perfect trial, and a second list, also of nonsense syllable-adjective pairs, for either 1, 5, 10, or 20 trials. They required Ss to recall both the first and second lists, and found that first list recall was inversely related to number of IL trials. Postman and Keppel (1967) used groups which received 1, 4, or 6 interpolated trials. A control group learned the original list and participated in a neutral interpolated activity, while the remaining groups learned a second list of words. All Ss studied the original list for four trials. Postman and Keppel (1967) found that recall of the original list on a terminal test of retention decreased progressively as a function of the number of interpolated trials. In contrast, recall of interpolated list words increased as a function of the number of interpolated trials. Their results suggest that interpolated learning produces effects in FR which are functionally the same as those noted in paired-associate situations.

Effect of Organization on Retroactive Inhibition in Free Recall

There have been a number of studies which have exam-

ined the effects of organization on RI in FR. Several of these studies (Shuell, 1968; Thompson, and Poling, 1969; Watts and Anderson, 1969; and Winograd, 1968) manipulated organization by presenting lists which either contained different words from the same categories on the two lists (S), or contained words from different categories on the two lists (D). The results of these studies suggest that there is more RI for S groups than for D groups. For example, Shuell (1968) presented Ss with two successive lists of 35 nouns representing seven conceptual categories. Half of the Ss (group S) received different words from the same categories on the two lists, while the other half of the Ss (group D) received words from different categories on the two lists. An outside control group participated in a neutral activity instead of learning a second list. By comparing recall of OL words for the S, D, and control groups, significant RI was indicated for both S and D groups, but the amount of RI was significantly greater for the S group. The RI in the D group resulted from the loss of whole categories, while the RI in the S group resulted from the loss of both whole categories and specific words.

An explanation of the above finding could be based on the suggestion by Tulving (1968), and McGovern (1964)

that RI should be obtained whenever there is a situation analogous to the A-B, A-C paradigm in paired-associate learning. If the "A" term represents the category, and the "B" and "C" terms represent the specific words, RI would be predicted for the S groups where "A" is constant over lists. Conversely, the D groups would be analogous to an A-B, C-D paradigm with different categories ("A", and "C") in the two lists, and less RI would be predicted.

Another way that organization has been manipulated in studies of RI in FR (Wood, 1970, 1971; Zavortnik and Keppel, 1968; and Royer, 1970) has been to give instructions to Ss about how they should organize words during study trials. These studies also found more RI when instructions were to organize both lists in the same fashion (S), rather than when instructions were to organize each list in a different fashion (D). For example, Wood (1970) conducted three experiments where the words to be recalled were held constant for all conditions. Organization was manipulated in Experiment I by instructing Ss how to organize the lists. The IL words were chosen so that they could be organized into the same or different categories as the OL words. The S group was instructed to use the same categories for both lists, while the D group was instructed to use different categories for the

two lists. Again, in Experiment II, IL words were chosen so that they could be categorized in the same, or in a different fashion than OL words. In this experiment Ss were told to sort the words according to the instructed categories. As in the first two experiments, Experiment III used the same lists for all conditions, but manipulated organization during presentation by using either alphabetical or categorical organization, and by instructing the S to use the appropriate organization during recall. All three of Wood's (1970) experiments resulted in more RI for the S groups than for the D groups.

In an attempt to demonstrate that the organizational variable does not always produce the same effect, Wood (1971) attempted to increase RI in groups which received different (D) instructions for organizing lists by requiring Ss to recall both lists during the final recall stage. He found that the S groups had a greater tendency to fail to recall specific items in categories, but D groups failed to recall entire categories. He, therefore, concluded that the effect of organization on RI depends on whether the conditions of the experiment favor the forgetting of whole categories (i.e., Ss are asked to recall many categories), or portions of categories (i.e., S are asked to recall a small number of large categories).

Support is given to Wood's notion by studies conducted by Strand (1970), and Tulving and Pstoka (1971), who demonstrated that the loss of whole categories represents a state in which memory units (categories) are available, but not accessible. This was demonstrated by the fact that providing Ss with category cues greatly reduced the amount of RI. A possible confounding factor in both of these studies (Strand, 1970; and Tulving and Pstoka, 1971) may be that t he words used were all high frequency associates of the categories according to category norms. When Ss were cued with category names, they may have free associated and come up with the correct items, rather than having actually recalled them.

Two other studies (Wood, 1969; and Lewis, 1971) investigated the effect of cueing on recall, but did not concern themselves with RI. In these studies lists were presented either blocked (B) by category, or randomly (R). Cueing consisted of giving one word from each category. Both studies found that cueing increased recall with the B, but not with the R presentation.

To summarize, the previous studies suggest that there will be more RI when there is similar organization between the two lists. This seems to hold regardless of whether

the organization involves category organization of the lists themselves, or the way in which Ss are told to organize the lists for recall. There is also some indication that RI can be affected by varying the number of categories used, or by cueing recall by giving the category name, or by giving one item from the category.

Blocked Versus Random Presentation

The present study involved using a B or R presentation. A number of studies have examined the effect of B versus R presentation in FR; however, the effect of the different modes of presentation on RI has not been investigated. An example of the work done is a study reported by Cofer, Bruce, and Reicher (1966), which showed that recall was augmented by B presentation. It should be noted, however, that this effect was only found for words which were high frequency associates of the category as indicated by category norms. Dallet (1964) also used B and R presentations in a study which varied the number of categories per list. He also found superior recall for the B presentation. In addition, he reported an interaction between mode of presentation and category size, with the greatest difference in recall between mode of presentation occurring in the condition of three words per category. In another study, Puff (1966) used lists containing ten words from each of three cate-

gories. The lists were constructed so that there were either 0, 9, 18, or 27 category repetitions in order of presentation. The 0 condition is therefore an R treatment, whereas the 27 condition is a B treatment. Puff found a linear relationship between the number of category repetitions and recall, again indicating that B presentation facilitates recall. In addition, Shuell (1969) reported an unpublished study by Cohen which used B and R presentationswith a 70 item list containing 20 categories. The results of this experiment indicated equivalent recall for both presentation modes. However, B groups recalled more words per category with fewer numbers of categories recalled. Shuell (1969) has suggested that this may indicate that "under certain conditions, B presentation may facilitate the coding or organization of the predetermined categories, while decreasing the likelihood that stable inter-category associations will be suggested." It may be that Cohen's findings are due to the large nurber of categories (20) which he used.

Shuell (1969) has suggested that B presentation is frequently considered to be more effective than R presentation for helping the S perceive the categorized nature of the list. The studies reported in this area support this assumption, and suggest that, in fact, B presentation

facilitates recall. There is, however, some question as to whether presentation mode interacts with word frequency association of the category (Cofer, Bruce, and Reicher, 1966). There is also some evidence to suggest an interaction between presentation mode and the number of categories per list (Dallet, 1964; Cohen 1969).

Clustering

The present study was also concerned in the way the order of words recalled is affected by presentation mode. Clustering, which was defined by Bousfield (1953) as "the sequential occurrence in free recall protocols of two items belonging to one of several categories represented in the stimulus list", was used to analyze organization of recall. Stimulus category repetition (SCR) was used as an index of clustering. The calculation of this index is described in Bousfield and Bousfield (1966).

There is some empirical evidence to suggest that clustering and recall are positively correlated (i.e., Mandler, 1967; and Tulving, 1968). However, Puff (1970), and Cofer, Bruce, and Reicher (1966) suggest that in some cases these measures may vary independently. In the study by Cofer, Bruce, and Reicher (1966) it was suggested that any relationship between recall and clustering may be dependent on a third factor, the frequency of word associates

to the category. By using B and R presentation modes they found facilitation of clustering by B presentation for both high and low frequency associate lists. However, they found facilitation of recall only with high frequency associate lists. A study which used categorized and non-categorized lists was conducted by Puff (1970). He found that Ss with high clustering scores who received categorized lists did no better than Ss who did not cluster during recall, but who received the same list. In contrast, Ss who received non-categorized lists recalled less than those who received categorized lists, regardless of whether or not those Ss who received the categorized list clustered their recall. This suggests that clustering was not the important variable. The determining factor seemed to be whether the lists were presented in a categorized or non-categorized fashion.

The above evidence suggests that B presentation leads to more clustering in recall than R presentation. There is also evidence (Bousfield, Berkowitz, and Whitmarch, 1959; Marshal, 1967; Robinson, 1966; and Shuell, 1968) that clustering increases over trials.

STATEMENT OF THE PROBLEM

Studies of the effect of organization on RI in FR have indicated that using the same organization in both OL and IL leads to more RI than using different organization. This is consistent with studies which used the A-B, A-C paired-associate paradigm. However, it seems premature to generalize this finding to all possible variations of organization. For example, it is suggested that blocked (B) presentation might produce less RI than random (R) presentation, even though both lists consist of words from the same categories.

The first suggestion that B presentation might produce less RI than R presentation was offered by Royer (personal communication). The basis for this suggestion was that he found little RI in groups receiving B presentation in a previous study. In addition, a pilot study was conducted by the author which supports this notion. In the pilot study 20 Ss were used: ten each in a B and an R condition. Categories in both the OL and IL lists were the same, while the specific words differed. The results indicated a significant difference with the B group exhibiting less RI. RI was measured by the number of words of the OL list which were recalled following study of the IL lists, with a smaller number of words.

recalled indicating more RI. See Table 1 for means and standard deviations of the number of words recalled on the criterion recall trial in the pilot study.

The present study investigated the effect of presenting both OL and IL lists in either B or R fashion. The lists consisted of 20 words: five from each of four categories. The B condition involved presentation of all words from the same category consecutively, while the R condition involved presentation of the words in a completely random order, regardless of category. It was hypothesized that groups receiving B lists would exhibit less RI than those receiving R lists even though the same categorical organization would be present.

In addition, a factor of OL-IL similarity was included. Same (S) groups received OL and IL lists containing different words of the same categories, while different (D) groups received OL and IL lists containing different categories. There was also a control (C) condition which worked on a neutral activity (arithmetic problems) during the IL phase of the experiment. It was hypothesized that RI would be greatest for the S groups and least for the C groups. This hypothesis is consistent with the previous literature on the effect of organization on RI in FR as reviewed above.

Table 1

Trials to criterion in OL, words recalled in IL, and OL words recalled in the criterion trial of the pilot study

	OL		IL		Criterion recall	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
B	2.40	0.70	72.80	3.99	13.40	2.84
R	3.00	0.94	63.00	2.98	10.40	2.37
$F(1,18)$ Values						
	$F=2.61$		$F=38.66$		$F=6.60$	
	$p>.05$		$p<.01$		$p<.05$	

Additional hypotheses were that there would be more clustering with B presentation than with R presentation, and that clustering would increase over trials. Support for these hypotheses is based on the literature already reviewed. In addition, support for this hypothesis was obtained in the pilot study. Clustering as measured by SCR was calculated for the first and final trials of OL, and for the first trial of IL, as well as for the criterion recall trial of OL words following study of IL words. The comparison between the pilot study groups is summarized in Table 2, and Figure 1 illustrates the change of clustering over trials.

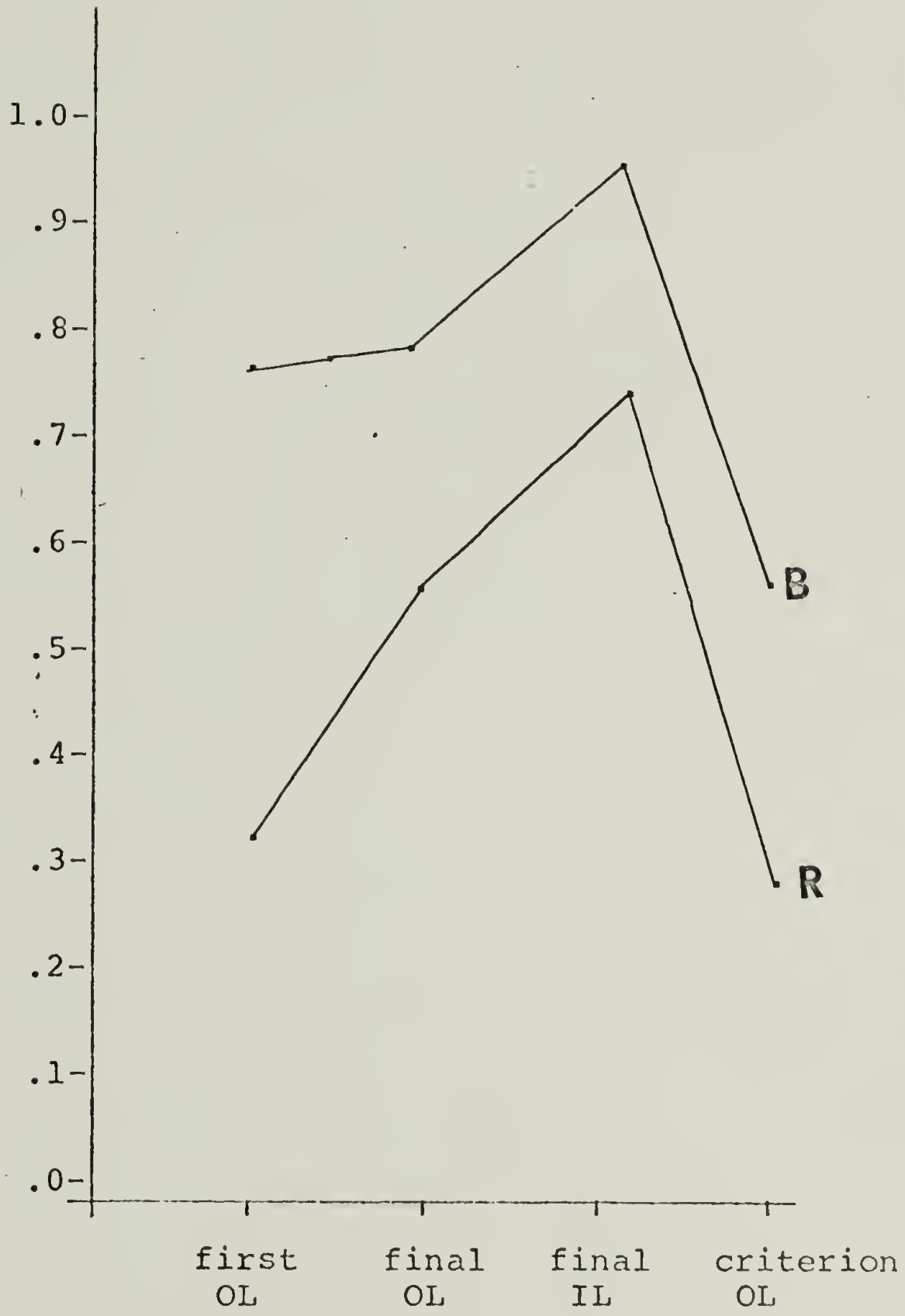
Table 2

SCR values in the pilot study

	First OL		Final OL		Final IL		Criterion	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
B	.75	.19	.78	.16	.96	.06	.58	.18
R	.31	.28	.58	.28	.72	.15	.28	.37
	F(1,18) Values							
	F=16.61		F=3.68		F=20.88		F=5.39	
	p<.01		p>.05		p<.01		p<.05	

Figure 1

SCR on selected trials
of pilot study



METHOD

Design

The experiment was a two by three completely randomized factorial design with an equal number of Ss (20) in each of the six cells. One independent variable was blocked (B), or random (R) list presentation, while the other independent variable was OL and IL lists with the same (S), or different (D) categories, or a control (C) group which did arithmetic problems during the IL phase. Table 3 outlines the design.

Subjects

There were a total of 120 Ss with 20 Ss randomly assigned to each of the six cells of the design. A restriction was that an equal number of males and females be assigned to each condition. All Ss were college students at the University of Massachusetts, who volunteered to participate in the experiment for extra credit in psychology courses.

Materials

All lists consisted of twenty words: five from each of four categories drawn from the less frequent associate half of the Battig and Montague (1969) category norms. See Table 4 for the lists of words which were used. Memory drum tapes were prepared with appropriate lists (S or

Table 3

Experimental design

Presentation mode	OL-IL Similarity: Same (S)	Different (D)	Control (C)
Blocked (B)	Group BS N=20 OL=IL B	Group BD N=20 OL≠IL B	Group BC N=20 IL problems B
Random (R)	Group RS N=20 OL=IL	Group RD N=20 OL≠IL	Group RC N=20 IL problems

Table 4

Lists of words

Categories	
countries	animals
crimes	sports
flowers	cloths
musical intruments	weapons

Lists			
Colombia	Bolivia	Beaver	Elk
Holland	Chile	Cheetah	Gazelle
Peru	Iceland	Llama	Jaguar
Rumania	Iraq	Raccoon	Panther
Yugoslavia	Hungary	Turtle	Skunk
Bigamy	Extortion	Boxing	Archery
Blackmail	Forgery	Diving	Boating
Felony	Fraud	Handball	Polo
Perjury	Homicide	Judo	Rugby
Treason	Manslaughter	Pool	Surfing
Aster	Buttercup	Acrilan	Calico
Camellia	Gladiola	Canvas	Chiffon
Marigold	Hyacinth	Crepe	Felt
Snapdragon	Magnolia	Gaberdine	Seersucker
Zinnia	Poppi	Lace	Taffeta
Accordian	Bassoon	Arrow	Bazooka
Bells	Cornet	Hatchet	Dagger
Bugle	Cymbols	Lance	Revolver
Fiddle	Harmonica	Shotgun	Scissors
Piccolo	Ukelele	Slingshot	Switchblade

D, and B or R) for each group. Three randomized orders of presentation were included on the tapes for each list. The B tape involved randomization of order of category presentation as well as randomization of words within each category. Group R tapes included presentations which were completely randomized regardless of category. The lists were counterbalanced for original or interpolated learning.

Procedure

The experimental session consisted of three phases: 1) acquisition of OL list; 2) acquisition of IL list; and 3) criterion recall of OL list. The acquisition phases (OL and IL) of the experiment consisted of alternating study and recall trials for all groups. During the study trials the words were presented on a memory drum at the rate of two seconds per word. Following each study trial there was a recall trial, which lasted for ninety seconds. During the recall trials the S was requested to orally recall all of the words he could remember from the list he just studied. He was told that he could recall words in any order he chose. Recall was recorded by the E.

Original learning. The original learning (OL) phase of the experiment continued until the S correctly recalled 18 of the 20 words. The following instructions were read to each S preceding OL:

> This is an experiment to determine how well you learn lists of words under specified conditions. I am going to show you a list of twenty words, at the rate of two seconds per word, on the memory drum in front of you. After you have seen the words, I will give you 90 seconds to recall as many words as you can. I will then show you the list of words again, and following this presentation there will be another 90 second recall period. This procedure will continue until you have recalled 18 of the 20 words in the list. You may recall the words in any order you choose. Are there any questions?

Interpolated learning. Following the correct recall of 18 words of the OL list, the interpolated learning (IL) phase of the experiment began. The C groups were given arithmetic problems to work on for 10 minutes (the same amount of time spent on IL by the other groups). The following instructions were read to each C S:

> We are interested in finding any relationship between arithmetic ability and the way you learn lists of words. I will therefore give you some arithmetic problems to work on now. Do as many of the problems as you can. I will tell you when your time is up. Are there any questions?

The other groups (S and D) studied and recalled an IL list in the same manner that they studied the OL list. For the IL list there were a total of four study and recall trials regardless of the number of words recalled on any trial. The following instructions were read to each S and D S preceding IL:

> I am now going to show you a second list of 20 words on the memory drum. I want you to learn this list in the same manner that you learned the first list. As

with the first list, you will be given 90 seconds to recall all the words that you can following each presentation trial. There will be a slight change in procedure for this list. Rather than stopping after the trial on which you recall 18 of the words, this time you will study and recall for a predetermined number of trials. I will tell you after the final trial is completed. Are there any questions?

Criterion recall. Following IL there was a criterion recall trial, to test the S's recall of OL words. The following instructions were read to each S preceding the criterion trial of OL words:

> Now what I want you to do is recall all the words you can from the (first)[1] list you studied. You can recall the words in any order you choose. Just tell me as many words as you can remember from the (first) list. You will have three minutes to recall the words. Are there any questions?

---

[1] words in parentheses were ommitted for C Ss.

RESULTS

Analyses of variance were used to detect significant sources of variance in the following dependent variables: trials to reach criterion in OL; words recalled in IL; OL words recalled in a criterion recall trial following IL; and clustering indices. The independent variables were the presentation factor, with the two levels being B or R presentation, and OL-IL similarity, with the levels being S, D, or C. In addition, sex of S was analyzed as a separate factor in hopes of reducing error variance. Since sex differences were not of interest in the present study, the sex main effect and interaction terms were pooled, and a single test was made, thus keeping type I error rate at a reasonable level (Anderson, 1968).

In addition to the analyses of variance, correlations were computed between dependent variables. Of particular interest were relations between OL learning and criterion recall, and relations between recall and clustering. Since the linear correlations were low, the scatter plots were examined in hopes of finding non-linear relations. The rationale for this approach was, in the case of recall and clustering, that clustering may have reached a ceiling and led to a curvilinear relationship between these two measures.

Table 5

Correlation matrix

	OL trials	IL words	OL criterion	SCR 1st OL	SCR final OL	SCR 1st IL	SCR final IL	SCR criterion
OL trials	1.00	-.43	-.05	-.33	-.14	-.25	-.29	-.10
IL words		1.00	.53	.25	.18	.47	.41	.27
OL criterion			1.00	.17	.00	.23	.20	.25
SCR 1st OL				1.00	.36	.38	.51	.30
SCR final OL					1.00	.29	.27	.43
SCR 1st OL						1.00	.63	.06
SCR final IL							1.00	.19
SCR criterion								1.00

OL Learning

Each S participated in OL until he correctly recalled 18 of the 20 words in a single trial. The mean number of trials to reach criterion and the standard deviation associated with each group can be seen in Table 6. As can be seen from the analysis of variance (Table 7) the main effect due to type of presentation (B versus R) was significant [$F(1,108)=11.43, p < .1$]. The nature of this effect was such that the R groups took more trials (mean=3.97) to reach criterion than the B groups (mean=3.23). No main effect was found for the OL-IL similarity factor ($F=1.12$). The means for the S, D, and C groups were, respectively 3.83, 3.53, and 3.45.

IL Learning

The C groups worked on arithmetic problems during IL, and were therefore not included in the analysis of IL learning. The Ss in the remaining groups all studied the IL lists for four trials and were compared as to the total number of words recalled in the four trials. Table 8 gives the means and standard deviations for this measure. The presentation factor was again significant [$F(1,72)=7.87$, p .01] as can be seen in Table 9. As in OL, the B presentation led to more recall than the R presentation. Summing over trials, the means for the B and R groups respectively

Table 6

Means and standard deviations for trials to criterion in OL

	S		D		C	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
B	3.45	1.46	3.30	1.26	2.95	0.94
B	4.20	1.51	3.75	1.07	3.95	1.23

Table 7

Analysis of variance for trials to criterion in OL

Source of Variance	df	MS	F	
Presentation (B versus R)	1	16.13	11.43	**
OL-IL similarity	2	1.58	1.12	
Presentation by OL-IL similarity	2	0.76	0.54	
Residual	6	2.27	1.61	
Subjects within cell	108	1.41		

** Significant at .01 level

Table 8

Means and standard deviations for words recalled in IL

	S		D	
	Mean	S.D.	Mean	S.D.
B	62.90	6.02	68.00	6.55
R	57.45	7.35	65.70	5.37

Table 9

Analysis of variance for words recalled in IL

Source of Variance	df	MS	F	
Presentation (B versus R)	1	296.45	7.87	**
OL-IL similarity	1	897.80	23.82	**
Presentation by OL-IL similarity	1	51.20	1.36	
Residual	4	91.32	2.43	
Subjects within cell	72	37.65		

** Significant at .01 level

were 65.45 and 61.60. There was also a main effect for IL learning due to OL-IL similarity [$F(1,72)=23.82$, $p<.01$]. The nature of this effect was that the D groups recalled more words than the S groups (means equal 66.88 and 60.18 respectively).

Criterion Recall of OL Words

Following IL all Ss were asked to recall as many words as they could from the first list they studied. Table 10 gives the means and standard deviations for the number of words recalled in this trial, and Table 11 gives the analysis of variance table. The presentation factor was not a significant source of variance for criterion recall of OL words ($F=1.3$). The main effect due to OL-IL similarity was significant [$F(2,108)=67.65$, $p<.01$]. The C groups recalled the most OL words (mean=17.73), followed by the D groups (mean=15.25), and the S groups recalled the fewest OL words (mean=9.45).

Multiple comparisons were used to determine which specific groups differed in criterion recall of OL words. A Dunnet test (Myers, 1966) was carried out to test the null hypothesis that experimental groups did not differ from the control groups. All four of the contrasts (BS versus BC, BD versus BC, RS versus RC, and RD versus RC) yielded d values larger than the criterion statistic required

Table 10

Means and standard deviations for OL words recalled in the criterion trial

	S		D		C	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
B	8.90	4.13	15.20	4.10	17.35	2.32
R	10.00	4.60	15.30	2.94	18.20	1.51

Table 11

Analysis of variance for OL words recalled in the criterion trial

Source of Variance	df	MS	F	
Presentation (B versus R)	1	14.01	1.30	
OL-IL similarity	2	728.81	67.65	**
Presentation by OL-IL similarity	2	2.73	0.25	
Residual	6	31.90	2.96	*
Subjects within cell	108	10.77		

* Significant at .05 level

** Significant at .01 level

for significance at the .01 level, [$\underline{d}(3,108)=2.60$, $\underline{p}<.01$], indicating that with probability of less than .01 a type I error occurred in at least one of the contrasts. In addition, a Bonferroni $\underline{t}$ test (Miller, 1966) was used to reject the null hypothesis that the S and D groups did not differ for this measure [$\underline{t}(108)=7.8$, $\underline{p}<.01$]. This test was based on a single *a priori* comparison. The fact that the Dunnet test was also carried out for this factor (however, separately for B and R groups) may lead to an error rate problem. However, the Bonferroni inequality ($EW \leq \Sigma\alpha_j$) suggests that the error rate for all of the above contrasts is less than or equal to the sum of the alpha levels used for the Dunnet and Bonferroni $\underline{t}$ tests (.02), and should, therefore, not be of concern.

Clustering Indices

SCR as described by Bousfield and Bousfield (1966) was computed for the first and final trials of OL, and IL, and for the criterion recall trial of OL words. Essentially, this index involves three values: PSCR (the maximum number of possible category repetitions based on the number of words recalled from each category); OSCR (the observed number of repetitions); and ESCR (the number of repetitions expected by chance based on the number of words recalled from each category). The final index used for comparison is computed

as follows:

SCR = OSCR - ESCR ÷ PSCR - ESCR.

This index has no negative limit, but a positive maximum of plus one, indicating recall totally by category. An index of zero indicates a number of repetitions equal to that expected by chance, or completely random order in recall.

The analysis of variance for clustering was computed with the same factors as the above analyses (D or R presentation, OL-IL similarity, and sex of S), with the additional within S variable of trials. Two incomplete analyses were used due to the fact that C groups did not have indices for IL trials. The first analysis (Table 12) included S, D, and C groups for the first and final OL trials and the criterion recall trial of OL words. The second analysis (Table 13) included S and D groups only, for the first and final trials of both OL and IL, and the criterion recall trial of OL words. Both analyses yielded a significant main effect for type of presentation [$F(1,108)=40.11$, $p<.01$] and [$F(1,72)=51.39$, $p<.01$] with the D groups clustering recall to a greater degree than the R groups on all trials. The means and standard deviations are given in Table 14. The main effect for OL-IL similarity was significant for the second analysis (experimental groups only over OL, IL, and criterion recall of OL words) [$F(1,72)=7.03$, $p<.01$], but not significant for

Table 12

Analysis of variance for SCR

(Experimental and control groups over first and final trials of OL and OL criterion trial)

Source of Variance	df	MS	F	
Presentation (B versus R)	1	5.32	40.10	**
OL-IL similarity	2	0.23	1.75	
Trials	2	3.28	48.67	**
Presentation by OL-IL similarity	2	0.27	2.01	
Presentation by trials	2	0.66	9.72	**
OL-IL similarity by trials	4	0.17	2.58	*
Presentation by OL-IL similarity by trials	4	0.15	2.22	
Residual	6	0.09	0.65	
Residual by trials	12	0.05	0.73	
Subjects within cell	108	0.13		
Subjects by trials within cell	216	0.07		

* Significant at .05 level

** Significant at .01 level

Table 13

Analysis of variance for SCR

(Experimental groups only over first and final trials of OL and IL and OL criterion trial)

Source of Variance	df	MS	F	
Presentation (B versus R)	1	7.06	51.39	**
OL-IL similarity	1	0.97	7.03	**
Trials	4	2.22	30.45	**
Presentation by OL-IL similarity	1	0.75	5.49	*
Presentation by trials	4	0.60	8.16	**
OL-IL similarity by trials	4	0.10	1.31	
Presentation by OL-IL similarity by trials	4	0.01	0.18	
Residual	4	0.02	0.12	
Residual by trials	16	0.06	0.87	
Subjects within cell	72	0.14		
Subjects by trials within cell	288	0.07		

* Significant at .05 level

** Significant at .01 level

Table 14

Means and standard deviations for SCR

		Trials								
		1st OL		final OL		1st IL		final IL		Criterion	
		Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
B	S	.70	.28	.88	.17	.77	.27	.93	.11	.80	.32
	D	.66	.39	.82	.22	.86	.22	.98	.04	.82	.26
	C	.71	.35	.87	.17	———	———	———	———	.92	.13
R	S	.14	.47	.65	.27	.27	.37	.68	.26	.59	.36
	D	.25	.44	.76	.20	.54	.39	.84	.22	.87	.13
	C	.45	.35	.58	.28	———	———	———	———	.72	.24

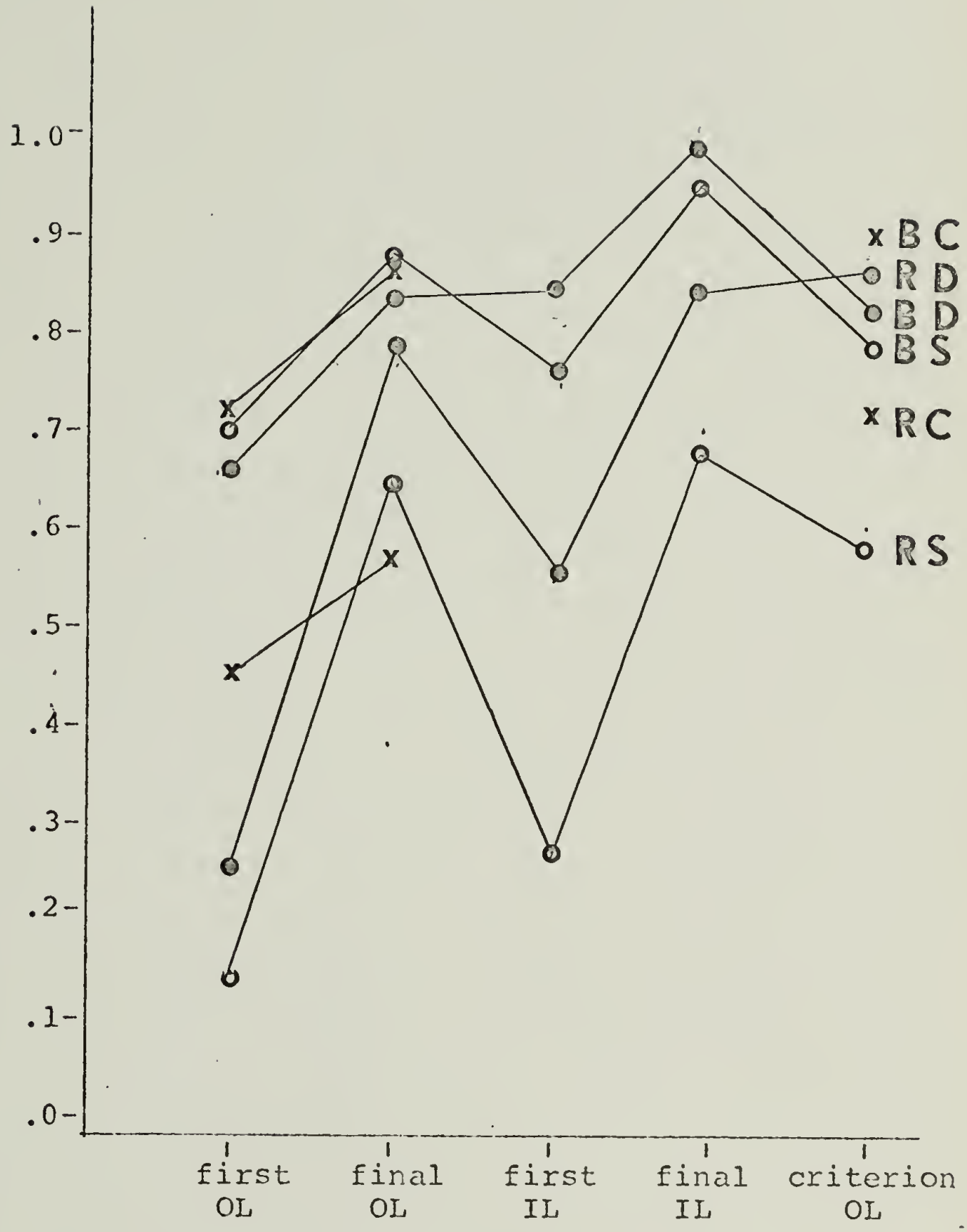
the first analysis (all groups over OL and criterion recall of OL words only) (F=1.75). This indicated that IL trials were the source of significance. The nature of this effect was that the D groups clustered to a greater extent in IL (mean=.81) than the S groups (mean=.66).

There was also a main effect due to trials for both analyses [$F(2,216)=48.67$, $p<.01$] and [$F(4,288)=30.45$, $p<.01$]. The nature of this main effect can be seen in Figure 2. That is, clustering increases from the first to the final trial of OL, and from the first to the final trial of IL. In addition, clustering is higher for IL trials than for OL trials.

In addition to the main effects mentioned above, there were a number of significant interactions. The interaction between presentation type and OL-IL similarity was significant [$F(1,72)=5.49$, $p<.05$] for the second analysis (experimental groups only over five trials). The nature of this interaction was that the B groups clustered to almost the same extent regardless of whether they had S or D lists (means equal .82 and .83), while the R groups clustered to a greater extent if they received D lists (mean=.65) than if they received S lists (mean=.46). Both analyses yielded a significant interaction between type of presentation and trials [$F(2,216)=9.72$, $p<.01$] and [$F(4,288)=8.16$, $p<.01$].

Figure 2

SCR on selected trials



While the R groups clustered to a lesser degree than the B groups overall, the increases in clustering over trials mentioned above were more pronounced for the R groups than for the B groups. Finally, the first analysis (all groups over first and final OL trials, and the criterion recall trial of OL words) yielded a significant interaction between OL-IL similarity and trials, [$F(4,216)=2.58$, $p<.05$]. Clustering for the criterion recall trial of OL words was greater than for the final trial for OL learning for D and C groups, but not for the S groups.

Correlations Between Dependent Variables

Correlations were computed between dependent variables as can be seen in the correlation matrix in Table 5. As would be expected, trials to criterion in OL learning and words recalled in IL learning were negatively correlated ($r=-.43$). That is, Ss who took longer to reach criterion in OL recalled fewer words in the four IL trials. On the other hand, trials to criterion in OL and words recalled in the criterion recall trial of OL words were not correlated ($r=-.05$), which indicates that RI was not affected by the number of trials it took the S to reach criterion in OL. The correlations between recall measures and clustering were low; this is especially the case for the OL criterion trial ($r=.27$). As would be expected, the intertrial cluster-

ing correlations were somewhat higher, although never greater than r=.63. Scatter plots were examined to determine whether non-linear relationships existed between some of these measures, however no such relationships were found.

DISCUSSION

The present study was concerned with the effect of list organization on RI. To assure that all groups learned the OL list equally well each S studied the list until he reached a criterion of 18 correctly recalled words on a single trial. The number of trials it took to reach criterion was used as a dependent measure of OL learning. The fact that the R groups took more trials to reach criterion than the B groups is consistent with previous literature which suggests that B presentation facilitates learning. The fact that S, D, and C groups did not differ on the number of trials it took to reach criterion in OL suggests that the groups were equivalent in OL learning. This finding was expected since there was no difference between the treatments during OL learning.

Since RI was of prime interest in the study, and since RI is influenced by the number of IL trials, all experimental Ss studied the IL words for the same number of trials (four). As in OL, the B presentation led to more recall than the R presentation, as measured by total number of words recalled in the four IL trials. There was also a main effect for IL learning due to OL-IL similarity with the D groups recalling more words than the S groups. This is an example of negative transfer, which is not common

in the free recall RI literature (c.f., Shuell, 1969; and Wood, 1970). The transfer can be interpreted via interference theory. That is, when the same categories were present in both lists (A-B, A-C paradigm) interference theory would predict lower recall than when different categories were present in the two lists (A-B, C-D paradigm).

The major focus of the present study was the experimental manipulation of RI. RI was measured by loss of OL words following IL that was present for the experimental groups, but not for the control groups. That is, the lower the OL criterion recall score, the more RI. As mentioned in the results section, all experimental groups differed significantly from the control groups as to criterion recall of OL words. Thus RI was present for all experimental groups. Empirical evidence, as reviewed in the introduction, suggests that similar organizations in OL and IL leads to more RI than dissimilar organizations. It is assumed that B presentation for both lists is an instance of more similar organization than R presentation for both lists (as indicated by clustering scores). However, it was hypothesized that R presentation would lead to more RI than B presentation. The fact that the B and R groups did not differ significantly on the measure of RI leads to failure to reject the null hypothesis that these treatments produce comparable effects.

However, the finding that B and R presentations produced comparable RI conflicts with the notion that all organizational manipulations affect RI in the same way (similar organizations leading to more RI than dissimilar organizations). The fact that there was no effect due to the presentation variable was disappointing considering the significant effect in the pilot study reported in the introduction. The BS group and the RS group in the present study are essentially identical to the groups in the pilot study. However, a comparison of these groups with the pilot study groups reveals obvious discrepancies in the outcomes of the two experiments. That is, in the present study the BS group recalled a mean of 8.90 OL words on the criterion trial, and the RS groups recalled a mean of 10.00 OL words, while in the pilot study the B group recalled a mean of 13.40 OL words on the criterion trial, and the R groups recalled a mean of 10.40 OL words. The inconsistency seems to lie in the difference between the B groups in the two studies, and may be attributed to either a type I error in the pilot study, or a type II error in the present study.

It was also predicted that S lists would lead to a more RI than D lists. That finding was supported, and replicates previous studied.

In addition to recall, the present study was concerned

with the effects of the independent variables on clustering. It was hypothesized that B presentation would lead to more clustering than R presentation, and this finding was supported.

The fact that D groups clustered to a greater extent in IL than S groups was not predicted, and is, at least to the present author, counter-intuitive. The categories should have been immediately apparent to S Ss, but not to D Ss, which would lead to a prediction in the opposite direction of the finding. There is some evidence that clustering and recall are highly related, (Kintch, 1970), and the fact that the S groups recalled fewer words in IL than D groups, and also clustered to a lesser extent is consistent with this notion.

The predictions concerning changes in clustering over trials were based on the pilot study, and the literature reviewed in the introduction. These predictions were supported. That is, clustering increased over OL trials and over IL trials. An obvious explanation for this finding is that the categories in the list became more apparent with successive trials. The fact that clustering is higher for IL than for OL is understandable if one assumes that the S is already aware of the categorical nature of the list, and in the case of S Ss, the categories themselves.

The interaction between presentation mode and OL-IL similarity for clustering can be attributed to the fact that the B groups clustered to a relatively high degree from the beginning, while the R groups had a larger range to increase clustering, and thus the S and D distinction became significant for the R groups.

The previous literature does not offer a definitive answer to the question of whether clustering and recall are correlated. The present study offers no support for the suggestion that they are, since the linear correlation coefficients between recall variables and clustering were low, and curvilinear relationships were not apparent from the scatter plots.

SUMMARY AND CONCLUSIONS

The present study was designed to investigate RI as a function of the organizational variable of B versus R presentation of categorized lists, and as a function of S versus D categories in the OL and IL lists. Previous studies on the effect of organization of lists on RI showed that similar organization between OL and IL led to more RI than dissimilar organization. The results of the present study concerning the use of S or D categories replicated previous studies, and supported the above notion. However, it was hypothesized that this rule of the effect of organization on RI might not hold for all types of organizational manipulations. Specifically, the present study looked at presentation of both lists in either a B or R fashion, the assumption being that the B presentation was a similar organization for OL and IL, while the R presentation was not. No significant differences were found between the B and R manipulations, although less RI was predicted for the B group than for the R group on the basis of a pilot study. While this finding does not indicate a case where similar organization is an asset (leads to less RI) it does suggest that similar organization does not necessarily result in more RI than dissimilar organization.

REFERENCES

Anderson, N. H. Partial analysis of high-way factorial designs. Behavioral Research Methodology and Instrumentation, 1968, 1, 2-7.

Barnes, J. M., & Underwood, B. J. "Fate" of first-list associations in transfer theory. Journal of Experimental Psychology, 1959, 58, 97-105.

Battig, W. F., & Montague, W. E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. Journal of Experimental Psychology Monograph, 1969, 80(3, Pt. 2).

Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates. Journal of General Psychology, 1953, 49, 229-240.

Bousfield, W. A., Berkowitz, H., & Whitmarch, G. A. Associative clustering in the recall of minimally meaningful geometric designs. Canadian Journal of Psychology, 1959, 13, 281-287.

Bousfield, A. K., & Bousfield, W. A. Measurement of clustering and sequential constancies in repeated free recall. Psychological Reports, 1966, 19, 935-942.

Briggs, G. E. Acquisition, extinction, and recovery functions in retroactive inhibition. Journal of Experimental Psychology, 1954, 47, 285-293.

Cofer, C. N., Bruce, D. R., & Reicher, G. M. Clustering in free recall as a function of certain methodological variations. Journal of Experimental Psychology, 1966, 71, 858-866.

Dallet, K. M. Number of categories and category information in free recall. Journal of Experimental Psychology, 1964, 68, 1-12.

Keppel, G. Retroactive and proactive inhibition. In T. R. Dixon, & D. L. Horton (Eds.), Verbal Behavior and General Behavior Theory. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

Kintsch, W. Learning, Memory, and Conceptual Processes. New York: John Wiley and Sons, Inc., 1970.

Lewis, M. Q. Categorized lists and cued recall. Journal of Experimental Psychology, 1971, 87, In press.

Mandler, G. Organization and memory. In K. W. Spence, & J. T. Spence (Eds.), The Psychology of Learning and Motivation. Vol. 1. New York: Academic Press, 1967.

Marshall, G. R. Stimulus characteristics contributing to organization in free recall. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 364-374.

McGeoch, J. A., & Irion, A. L. The Psychology of Human Learning. New York: Longmans, Green, 1952.

McGovern, J. B. Extinction of association in four trans-

fer paradigms. Psychological Monograph, 1964, 78, (16, Whole No. 593).

Melton, A. S. Comments on Professor Postman's paper. In C. N. Cofer (Ed.), Verbal Learning and Verbal Behavior. New York: McGraw, 1961.

Melton, A. W., & Irwin, J. M. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. American Journal of Psychology, 1940, 53, 173-203.

Miller, R. G., Jr. Simultaneous Statistical Inference. New York: McGraw-Hill, 1966.

Myers, J. L. Fundamentals of Experimental Design. Boston: Allyn and Bacon, 1966.

Postman, L. The present status of interference theory. In C. N. Cofer (Ed.), Verbal Learning and Verbal Behavior. New York: McGraw-Hill, 1961.

Postman, L., & Keppel, G. Retroactive inhibition in free recall. Journal of Experimental Psychology. 1967, 74, 203-211.

Puff, C. R. Clustering as a function of sequential organization of stimulus word lists. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 503-506.

Robinson, J. A. Category clustering in free recall. Journal of Experimental Psychology, 1966, 62, 279-

285.

Royer, J. M. Retroactive inhibition in free recall: A storage location hypothesis. Unpublished doctoral dissertation, University of Illinois, 1970.

Scheffe, H. A. The Analysis of Variance. New York: John Wiley and Sons, Inc., 1970.

Shuell, T. J. Retroactive inhibition in free recall learning of categorized lists. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 797-805.

Shuell, T. J. Clustering and organization in free recall. Psychological Bulletin, 1969, 72, 353-374.

Slamecka, N. J., & Ceraso, J. Retroactive and proactive inhibition of verbal learning. Psychological Bulletin, 1960, 57, 449-475.

Strand, B. Z. A further investigation of retroactive inhibition in categorized free recall. Journal of Experimental Psychology, 1971, 87, In press.

Swenson, E. J. Retroactive inhibition: A review of the literature. Minnesota Studies on Education, 1941, 1.

Thompson, C. P., & Poling, D. R. Free recall of successive lists: Effect of category size and category repetition. Psychonomic Science, 1969, 16, 201-202.

Tulving, E. Theoretical issues in free recall. In T. R. Dixon, & D. L. Horton (Eds.), Verbal Behavior

and General Behavior Theory. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

Tulving, E., & Pstoka, J. Retroactive inhibition in free recall: Inaccessibility of information available in memory store. Journal of Experimental Psychology, 1971, 87, 1-8.

Tulving, E., & Thornton, G. B. Interaction between proaction and retroaction in short-term retention. Canadian Journal of Psychology, 1959, 13, 155-165.

Underwood, B. J. Retroactive and proactive inhibition after five and forty-eight hours. Journal of Experimental Psychology, 1948a, 38, 19-38.

Underwood, B. J. "Spontaneous Recovery" of verbal associations. Journal of Experimental Psychology, 1948b, 38, 429-439.

Watts, G. H., & Anderson, R. C. Retroactive inhibition in free recall as a function of first- and second-list organization. Journal of Experimental Psychology, 1969, 81, 595-597.

Winograd, E. List differentiation, recall and category similarity. Journal of Experimental Psychology, 1968, 78, 510-515.

Wood, G. Retrieval cues and the accessibility of higher-order memory units in multi-trial free recall. Journal of Verbal Learning and Verbal Behavior,

1969, 8, 282-289

Wood, G. Organization, retroactive inhibition, and free recall learning. Journal of Verbal Learning and Verbal Behavior, 1971, 10, In press.

Zavortnik, B., & Keppel, G. Retroactive inhibition in free recall learning with alphabetical cues. Journal of Experimental Psychology, 1968, 78, 617-624.